IAI Journals.ASM.org

Analysis of *Helicobacter pylori cagA* Promoter Elements Required for Salt-Induced Upregulation of CagA Expression

John T. Loh,[a] David B. Friedman,[b] M. Blanca Piazuelo,[c] Luis E. Bravo,[d] Keith T. Wilson,[c,e,f] Richard M. Peek, Jr.,[c] Pelayo Correa,[c] and Timothy L. Cover[a,e,f]

Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA[a]; Proteomics Laboratory, Mass Spectrometry Research Center, Vanderbilt University School of Medicine, Nashville, Tennessee, USA[b]; Division of Gastroenterology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA[c]; Department of Pathology, Universidad del Valle School of Medicine, Cali, Colombia[d]; Department of Pathology, Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, Tennessee, USA[e]; and Veterans Affairs Tennessee Valley Healthcare System, Nashville, Tennessee, USA[f]

Helicobacter pylori **infection and consumption of a high-salt diet are each associated with an increased risk for the development of gastric cancer. To investigate potential synergism between these factors, we used a global proteomic approach to analyze *H. pylori* strains cultured in media containing varying salt concentrations. Among the differentially expressed proteins identified, CagA exhibited the greatest increase in expression in response to high salt concentrations. Analysis of 36 *H. pylori* strains isolated from patients in two regions of Colombia with differing incidences of gastric cancer revealed marked differences among strains in salt-responsive CagA expression. Sequence analysis of the *cagA* promoter region in these strains revealed a DNA motif (TAATGA) that was present in either one or two copies. Salt-induced upregulation of CagA expression was detected more commonly in strains containing two copies of the TAATGA motif than in strains containing one copy. Mutagenesis experiments confirmed that two copies of the TAATGA motif are required for salt-induced upregulation of CagA expression. In summary, there is considerable heterogeneity among *H. pylori* strains in salt-regulated CagA expression, and these differences are attributable to variation in a specific DNA motif upstream of the *cagA* transcriptional start site.**

Helicobacter pylori is a Gram-negative bacterium that persistently colonizes the human gastric mucosa. *H. pylori* colonization results in gastritis and is associated with an increased risk of gastric cancer and peptic ulceration. The clinical outcome of *H. pylori* infection is determined by a combination of bacterial, host, and environmental factors (1, 2, 9, 31, 35).

One of the most important *H. pylori* virulence determinants is the CagA protein. CagA is delivered into host gastric epithelial cells by a type IV secretion system and is subsequently phosphorylated by host cell tyrosine kinases at conserved EPIYA motifs within the CagA protein (3, 16, 21, 37). Within gastric epithelial cells, CagA can interact with multiple host cell targets, leading to numerous alterations in cell signaling and morphology (3, 16, 21, 37). Epidemiological studies have demonstrated a higher risk of gastric cancer in persons infected with *cagA*-positive *H. pylori* strains than among persons infected with *cagA*-negative strains (6, 7, 11, 20, 29, 30). In addition, strains expressing CagA proteins with a high copy number of EPIYA motifs have been associated with a higher risk of disease than strains expressing CagA proteins with fewer EPIYA motifs (4, 18, 26). Recently, an analysis of *H. pylori* strains cultured from patients living in regions of Colombia with disparate risks for gastric cancer revealed differences in the levels of CagA expression (23). *H. pylori* strains expressing higher levels of CagA were associated with more advanced precancerous gastric lesions, compared to the histologic abnormalities found in persons infected with strains expressing low levels of CagA (23).

A high dietary intake of salt has been associated with an increased risk for the development of gastric cancer (reviewed in reference 38), and synergistic effects of salt and *H. pylori* in the development of gastric cancer have been reported in animal models (28, 36). However, the mechanism by which high salt intake contributes to the development of gastric carcinoma is poorly understood. One possibility is that a high-salt diet may have direct effects on gastric tissue that predispose to the development of gastric cancer. Another possibility is that a high-salt diet may cause alterations in *H. pylori*.

Several previous studies have analyzed the effects of differing salt concentrations on *H. pylori*. In response to high sodium chloride (NaCl) concentrations, *H. pylori* cells change from a typical spiral shape to a more elongated shape and form chains (13, 14). In addition, differences in salt concentration lead to alterations in *H. pylori* gene expression (13, 25). One study utilized array methodology to identify multiple *H. pylori* genes that were regulated in response to changes in salt concentration; most of the changes observed were not validated by using other methods, but upregulation of *cagA* expression in response to increased levels of NaCl was validated by the use of transcriptional reporters and Western blotting (25). Other studies reported that the expression of *vacA*, *amiE*, and *pfr* was regulated in response to differences in salt concentration (13, 14), but these studies failed to detect any effect of salt on CagA expression. Thus, there has been relatively little consensus among previous studies in the identification of *H. pylori*

Received 4 March 2012 **Returned for modification** 5 April 2012
Accepted 8 June 2012
Published ahead of print 18 June 2012
Editor: S. R. Blanke
Address correspondence to Timothy L. Cover, timothy.l.cover@vanderbilt.edu.
Supplemental material for this article may be found at http://iai.asm.org/.

doi:10.1128/IAI.00232-12

genes that are regulated in response to differences in salt concentrations.

In the current study, we used a discovery-based proteomic profiling approach (2-dimensional difference gel electrophoresis [2D-DIGE]) as a nonbiased comprehensive approach with which to evaluate the effects of salt on protein expression in *H. pylori*. We confirm that CagA expression is significantly upregulated when *H. pylori* is cultured in a medium containing elevated salt concentrations, and we detect differential expression of multiple other proteins. We then examine salt-regulated CagA expression in *H. pylori* strains cultured from patients in two regions of Colombia with differing incidences of gastric cancer. We report that there is a considerable difference in the capacity of these strains to modulate CagA expression in response to increased salt concentrations in the bacterial culture medium. We analyze nucleotide sequences upstream of the *cagA* transcriptional start site, detect a motif that is present in different copy numbers, and report that the presence of two copies of this motif is associated with salt-induced upregulation of CagA expression. Finally, we show through a series of mutagenesis experiments that these motifs play an important role in salt-induced upregulation of CagA expression.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The *H. pylori* strains used in this study were isolated from gastric antral mucosa biopsy samples from individuals in the state of Nariño, Colombia (10, 23). The 36 strains analyzed in this study (17 from a region of low cancer risk and 19 from a region of high cancer risk) were all *cagA*-positive, contained type s1m1 *vacA*, and were randomly selected from a previously characterized set of Colombian strains (23). *H. pylori* strains were grown in room air supplemented with 5% CO_2 at 37°C. Bacterial cultures were routinely passaged every 2 days on Trypticase soy agar plates containing 5% sheep blood. In addition, *H. pylori* strains were grown in sulfite-free *Brucella* broth medium supplemented with 5% fetal bovine serum (BB-FBS) (25). The BB-FBS medium routinely used for growing *H. pylori* was prepared by addition of 0.5% NaCl (BB-FBS-0.5%). In the current study, the salt concentration of the BB-FBS was altered by adding either 0.25% NaCl (43 mmol/liter) or 1.25% NaCl (215 mmol/liter). When necessary, BB-FBS agar plates were supplemented with metronidazole (7.5 μg/ml) or chloramphenicol (5 μg/ml). *Escherichia coli* strains were grown in Luria-Bertani (LB) medium containing ampicillin (50 μg/ml) or chloramphenicol (25 μg/ml).

Human subjects. The *H. pylori* strains used in this study were isolated from 36 male subjects who underwent gastrointestinal tract endoscopy in two public hospitals in the State of Nariño, Colombia (10, 23). The hospitals are located in two cities with contrasting gastric cancer risks: Túquerres in the Andes Mountains, where the gastric cancer incidence is high (approximately 150 cases per 100,000), and Tumaco on the coast, where the gastric cancer incidence is low (approximately 6 cases per 100,000). The individuals included in this study are part of a larger series of subjects; the characteristics of the populations and exclusion criteria have been described previously (10).

Analysis of histopathology. A global gastric histologic diagnosis for each subject was determined, as described previously (10, 23). Global diagnoses were assigned on an ordinal scale from 1 to 6, as follows: 1, mild to moderate nonatrophic gastritis; 2, severe nonatrophic gastritis; 3, multifocal atrophic gastritis without intestinal metaplasia (MAG); 4, intestinal metaplasia (IM); 5, dysplasia; and 6, adenocarcinoma. A previously validated histopathology scoring system, which takes into account the extent and degree of atrophic, metaplastic, and dysplastic changes, was also used to quantify differences in morphological variables within each global diagnosis category (10, 23).

Cultivation of bacteria for proteomic analyses. *H. pylori* strains 26695 and 7.13 were each grown overnight in BB-FBS-0.5% to an optical density at 600 nm (OD_{600}) of about 0.5 and then subcultured (1:10) into fresh BB-FBS-0.25%, BB-FBS-0.5%, or BB-FBS-1.25% for 15 h. Bacterial cultures were harvested, and bacteria were lysed for 30 min with lysis buffer {7 M urea, 2 M thiourea, 4% 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS)}, as described previously (22).

DIGE-MS. Quadruplicate cultures of *H. pylori* were grown in BB-FBS-0.25%, BB-FBS-0.5%, or BB-FBS-1.25%. The *N*-hydroxysuccinimide (NHS)-ester dyes Cy2, Cy3, and Cy5 were used for minimal labeling, and lysates were resolved using difference gel electrophoresis (DIGE) gels and the mixed-sample internal standard methodology, as described previously (22). Each DIGE gel contained 100 μg of each of two samples along with a 100-μg aliquot of the Cy2-labeled standard (300 μg total per gel). Both pH 4 to 7 and pH 3 to 11 ranges were used for isoelectric focusing. DIGE analysis was performed with the DeCyder (version 6.5) suite of software tools (GE Healthcare), with each individual protein spot feature from a Cy3- or Cy5-labeled sample directly quantified relative to the Cy2-labeled signal from the pooled-sample internal standard corresponding to the same spot feature. A total of 977 protein features were matched across the gel series for purposes of quantification; from these, we selected the subset of proteins showing statistically significant changes for subsequent protein identification. The statistical significance of observed changes in abundance or charge-altering posttranslational modification was calculated using Student's *t* test and analysis of variance (ANOVA). Unsupervised principal component analysis (PCA) was performed using the DeCyder Extended Data Analysis (EDA) module. Differentially expressed proteins were robotically excised and were digested in-gel with modified porcine trypsin protease (Trypsin Gold; Promega, Madison, WI). The resulting peptides were analyzed by liquid chromatography-tandem mass spectrometry using an LTQ linear ion trap tandem mass spectrometer (MS) equipped with a Micro AS autosampler, a Surveyor high-performance liquid chromatography (HPLC) pump, a nanospray source, and Xcalibur 2.0 instrument control (Thermo Scientific, San Jose, CA). Candidate proteins were identified by searching the MS data against a database of *H. pylori* sequences extracted from the UniProtKB database (www.uniprot.org), which was concatenated with the database sequences in reverse so as to enable false-discovery-rate calculations. The database also included sequences of common laboratory contaminants. Searches were performed using the Sequest database search algorithm, allowing for cysteine carbamidomethylation and partial methionine oxidation.

Real-time PCR. Total RNA was isolated from *H. pylori* using TRIzol reagent (Gibco) as described previously (23). Following digestion with RQ1 RNase-free DNase (Promega), the RNA samples were subjected to a cleanup step using RNeasy columns (Qiagen). One hundred nanograms of purified RNA was then used for first-strand cDNA synthesis (Bio-Rad). Control cDNA reactions were carried out in the absence of reverse transcriptase to ensure the absence of contaminating genomic DNA from the RNA preparation. The cDNA and control reaction mixtures were diluted 1:20 and were used in real-time PCRs. Real-time PCR was performed using an ABI Real Time PCR machine, with SYBR green as the fluorochrome. Transcript abundance was calculated using the $\Delta\Delta C_T$ method, with each transcript signal normalized to the abundance of the 16S rRNA internal control. The normalized transcript signals obtained for *H. pylori* grown in BB-FBS-1.25% were then compared to the normalized transcript signals obtained for *H. pylori* grown in BB-FBS-0.5%. The primers used for real-time analysis are listed in Table 1.

Nucleotide sequence analysis of *cagA* promoter regions. The nucleotide sequences of a region upstream of the *cagA* ATG initiation site in 36 Colombian *H. pylori* strains were determined as described previously (23). These sequence data comprise about ~530 bp upstream of the ATG initiation site and include the *cagA* promoter region.

Introduction of *cagA* promoter regions from Colombian strains into *H. pylori* 26695. *cagA* promoter sequences from Colombian strains were introduced into *H. pylori* strain 26695 *cagA*::*catrdx*-9 using a previously described protocol (23). Strain 26695 *cagA*::*catrdx*-9 contains a *cat-rdxA* cassette inserted immediately downstream of the ATG initiation

TABLE 1 List of primers used in this study

Primer name	DNA sequence
Inverse PCR mutagenesis primers	
CagAprom1	5′-TTGTGCAAGAAATTCCATGAAA-3′
CagAprom2	5′-AATGGATCCGGCTCATTTCTTATTTCTTGTTC-3′
CagAprom5	5′-AATGGATCCCGCCTACAACGATCGGGCTT-3′
CagAprom6	5′-AATGGATCCGCTTGATATTGTTGTATAATAAGAATGTTC-3′
CagAprom7	5′-AATGGATCCAAGAATGTTCAAAGACACGAATTGAC-3′
CagApromR1	5′-AATGGATCCGGAAGCAAGATGAATGCAAGG-3′
CagApromR2	5′-AATGGATCCATAAGAACAAGAAATAAGAAATGAGCC-3′
CagApromR3	5′-AATGGATCCGTCATACTTACCATAGTATGC-3′
CagApromR4	5′-AATGGATCCCGTTTGTGATCCACATTGGC-3′
CagApromR5	5′-AATGGATCCAAGCCCGATCGTTGTAGGCG-3′
Real-time PCR primers	
16S rRNA F	5′-GGAGTACGGTCGCAAGATTAAA-3′
16SrRNA R	5′-CTAGCGGATTCTCTCAATGTCAA-3′
cagA-F	5′-GAGTCATAATGGCATAGAACCTGAA-3′
cagA-R	5′-TTGTGCAAGAAATTCCATGAAA-3′
hopQ-F	5′-TTGAGATCGGTGTTAGGGCTATG-3′
hopQ-R	5′-TGTGCTCCCACCGCAATT-3′
ureA-F	5′-GAAGACATCACTATCAACGAAGGCAA-3′
ureA-R	5′-GTTACCGCCAATGTCAATCAA-3′
vacA-F	5′-ACAACAAACACACCGCAAAA-3′
vacA-R	5′-CCTGAGACCGTTCCTACAGC-3′
HP0014-F	5′-GCCTAATTTTTTCATGGCGTTTAA-3′
HP0014-R	5′-TTGTTACACACCCCCCCATAA-3′
HP0027-F	5′-ACCCCATTATCCCCTTCATTG-3′
HP0027-R	5′-GATAGCATTTTTCGCCCACAA-3′
HP0310-F	5′-CCACGGCTTTTGAAATTGTT-3′
HP0310-R	5′-TGAGCGAGTGGTCGTATTTG-3′
HP0516-F	5′-TAAAACTGAAATCGCAAGACGAATAG-3′
HP0516-R	5′-ATCGCGCCCCACAAAAC-3′
HP0630-F	5′-GGGAGCTCTGGAGGGAAACT-3′
HP0630-R	5′-CGTCGCATCAGCGCTAAAG-3′
HP0642-F	5′-GGGAGATTAGCCCCTTCTTCA-3′
HP0642-R	5′-TCGCGCAAGATAAATGACAAA-3′
HP0688-F	5′-CGCGCCTGATAAGCTTCAAT-3′
HP0688-R	5′-CCAAATGGAACTGCTCGCTATA-3′
HP0979-F	5′-AAGGCTTTGCGTTAATGGGTATT-3′
HP0979-R	5′-AATCAGGGTGGTGCTCAAAAA-3′
HP1588-F	5′-GCTGTCATTGTTGCGAATGC-3′
HP1588-R	5′-TGCAATATCAATCGCTGTCCATA-3′

site of the *cagA* gene; the *cat-rdxA* cassette confers resistance to chloramphenicol mediated by the chloramphenicol acetyltransferase (*cat*) gene from *Campylobacter coli*, and susceptibility to metronidazole is mediated by an intact *rdxA* gene (HP0954) from *H. pylori* 26695. Briefly, DNA fragments encompassing approximately 0.5 kb upstream and 0.7 kb downstream from the *cagA* transcription start site were PCR amplified from Colombian strains using primers 5′-GCAAAAACAAACCCAAGCTGA-3′ and 5′-ACCAGTAGGCCCTCCATTTTTTTC-3′, and the amplified products were cloned into the pGEM-T Easy (Promega) vector. The resultant plasmids, which are unable to replicate in *H. pylori*, were used to transform *H. pylori* strain 26695 *cagA*::*catrdx*-9 (23, 32). Metronidazole-resistant transformants (arising due to homologous recombination) were selected by culturing on BB-FBS agar plates containing metronidazole. Metronidazole-resistant and chloramphenicol-susceptible colonies arising from the transformation represent recombination events in which the *cat-rdxA* cassette is removed and replaced with transformed sequences. All *H. pylori* transformants were subsequently screened for CagA expression by Western blotting, and CagA-expressing strains were sequenced to verify that sequences in the 26695 *cagA* promoter region had been replaced with the corresponding sequences from the Colombian strains.

Deletion and site-directed mutagenesis of the *H. pylori* 26695 *cagA* promoter region. To introduce deletion mutations into the promoter region of *cagA*, a 1.2-kb fragment of *cagA* encompassing 0.6 kb of DNA upstream and 0.6 kb of DNA downstream of the transcription initiation site of *cagA* was PCR amplified from *H. pylori* strain 26695 with primers 5′-AACCGACGCTTTTGTTTGGTAA-3′ and 5′-GATAGGGGGTTGTATGATATTTTC-3′ and was cloned into the plasmid vector pGEM-T Easy to yield pCagAprom18. Inverse PCR was used to introduce deletion mutations upstream of the *cagA* transcriptional start site into this plasmid. The primers used for the inverse PCRs are shown in Table 1. A BamHI site is contained within the primers to facilitate the ligation of inverse PCR products. The presence of the expected deletions was confirmed by DNA sequencing, and the mutated plasmids were used to transform *H. pylori* strain 26695 *cagA*::*catrdx*-9. Mutants were then selected by screening for metronidazole resistance and sensitivity to chloramphenicol as described above.

To introduce site-specific mutations, we used Quick Change mutagenesis (Agilent Technologies) with plasmid pCagAprom18 as a template. The introduction of desired mutations into plasmids was confirmed

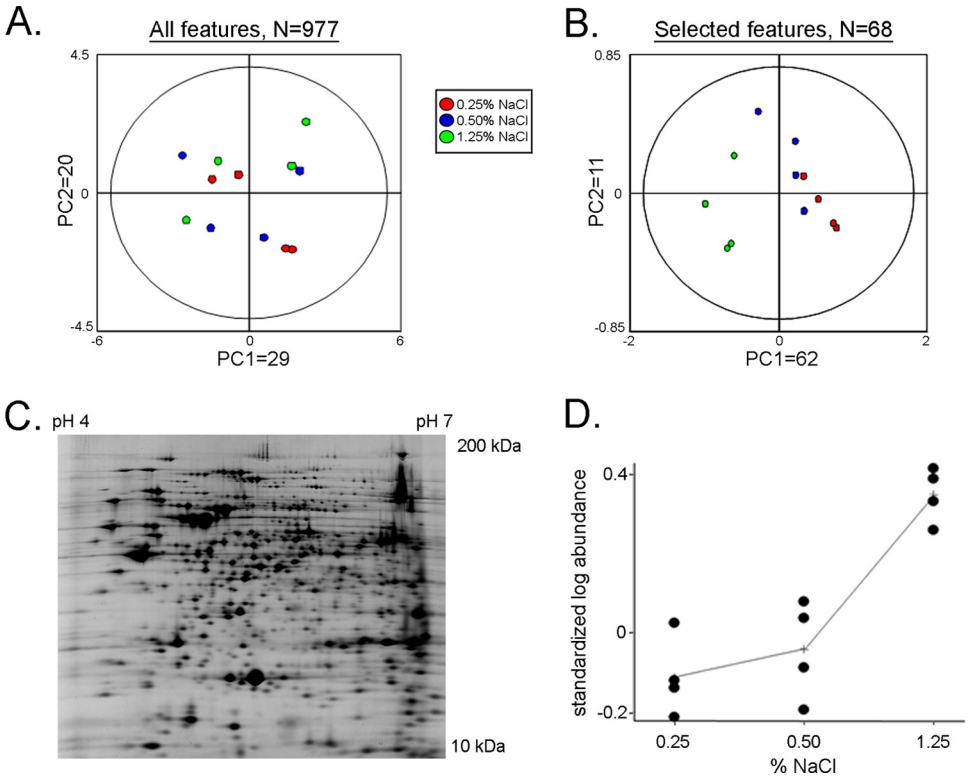

FIG 1 Proteomic analysis of *H. pylori* strain 26695 cultured in media containing varying NaCl concentrations. *H. pylori* 26695 was cultured for 15 h in media containing varying NaCl concentrations (BB-FBS-0.25%, BB-FBS-0.5%, or BB-FBS-1.25%). Cell lysates were generated and were subjected to 2D-DIGE analysis as described in Materials and Methods. (A) Unsupervised principal component analysis was used to assess differences in protein expression patterns among 12 samples (quadruplicate samples of *H. pylori* 26695 grown in BB-FBS-0.25%, BB-FBS-0.5%, or BB-FBS-1.25%). In total, 977 resolved protein spot features were matched across gels and were quantified. The first principal component (PC1), comprising 29% of the variation, did not distinguish between the three growth states (low, normal, and high salt concentrations). (B) Principal component analyses of 68 protein features that were altered in expression in any one of the three groups of samples relative to the others (based on *P* values of <0.05 by ANOVA). PC1 comprises >62% of the variation among the 68 features in this subset and distinguishes between bacteria grown under high-salt conditions and those grown under low- or medium-salt conditions. (C) Sypro ruby stain of a representative 2D-DIGE gel (pH 4 to 7) containing three differentially labeled samples as described in Materials and Methods. (D) Graph comparing CagA abundance between the three growth states (low, normal, and high salt concentrations). Standardized log abundance (*y* axis) was calculated relative to the signals from the Cy2-labeled internal standard described in Materials and Methods.

by DNA sequencing. Mutated plasmids were then used to transform *H. pylori* strain 26695 *cagA*::*catrdx*-9 as described above.

Western blot analysis. To analyze CagA expression in *H. pylori* strains, the bacterial strains were inoculated into BB-FBS-0.5% and were grown overnight. Cultures were then inoculated at an initial OD_{600} of ~0.1 in fresh BB-FBS-0.5% or BB-FBS-1.1%. Broth cultures were grown for 15 h and were harvested; the OD_{600} of the *H. pylori* cultures at the time of harvest was ~0.5 to 0.6. *H. pylori* cells were lysed using NP-40 lysis buffer, and each sample (5 μg of protein) was subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot analysis. CagA expression was analyzed using a 1:10,000 dilution of a CagA-specific polyclonal antiserum generated against amino acids 1 to 300 (Santa Cruz Biotechnology) and a horseradish peroxidase (HRP)-conjugated anti-rabbit antiserum (1:6,000; Promega). Immunoreactive bands were visualized by chemiluminescence (Roche). Following immunoblotting to detect CagA, blots were treated with Restore Western blot stripping buffer (Thermo) and were reprobed with either a 1:10,000 dilution of an antiserum generated against UreA (Santa Cruz Biotechnology) or a 1:10,000 dilution of an antiserum generated against soluble *H. pylori* proteins (anti-Hp). Densitometry was performed using ImageJ software (NIH) to quantify CagA and UreA signal intensities. Following subtraction of background signal levels, CagA signal intensities were normalized to the UreA signal for each sample (i.e., CagA/UreA). To quantify levels of salt-induced CagA expression, the normalized CagA signal of strains grown in BB-FBS-1.1% was divided by the normalized CagA signal obtained for the same strain grown in BB-FBS-0.5%, resulting in a CagA induction value. For each strain, the effects of NaCl on CagA expression were analyzed in a minimum of 3 independent experiments, and an average CagA induction value (comparing growth in 1.1% NaCl to growth in 0.5% NaCl) was calculated.

Nucleotide sequence accession numbers. The nucleotide sequences of a region upstream of the *cagA* ATG initiation site in 36 Colombian *H. pylori* strains, determined in this study, have been deposited in GenBank under accession numbers JX047554 to JX047589.

RESULTS

Analysis of *H. pylori* protein expression in response to varying salt concentrations. To investigate the effect of salt on *H. pylori* protein expression, wild-type strain 26695 was cultured in media containing varying NaCl concentrations: BB-FBS-0.25% (designated the low-salt medium), BB-FBS-0.5% (designated the normal medium), or BB-FBS-1.25% (designated the high-salt medium). As described previously, the growth rates of *H. pylori* in these different media were similar (25), and bacteria were harvested from all of the cultures at an OD_{600} of about 0.5. The bacteria were then analyzed in a global proteomic profiling experiment (2D-DIGE), as described in Materials and Methods. In total, 977 resolved protein spot features

TABLE 2 *H. pylori* proteins that are differentially expressed in response to environmental salt concentration

Protein[a]	Gene[b]	Function[c]	Value for bacteria grown in media containing different NaCl concns[d]					
			Strain 26695				Strain 7.13 (1.25% vs 0.5% NaCl)	
			1.25% vs 0.25% NaCl		1.25% vs 0.5% NaCl			
			Avg vol ratio	*P*	Avg vol ratio	*P*	Avg vol ratio	*P*
Downregulated[e]								
HspB	HP0010	GroEL					0.79	0.021
MetB	HP0106	Cystathionine gamma synthetase	0.68	0.031	0.8	0.120		
NeuB	HP0178	Sialic acid synthase	0.71	0.031	0.78	0.009		
MetK	HP0197	*S*-Adenosylmethione synthetase	0.75	0.022	0.85	0.120		
NifU	HP0221	FeS cluster assembly	0.75	0.006	0.75	0.009		
VacA paralog	HP0289	Unknown					0.65	0.034
DsbC	HP0377	Thioredoxin fold type protein	0.41	<0.001	0.54	0.007		
HslU	HP0516	Heat shock protein	0.75	0.017	0.75	0.015		
FlaA	HP0601	Motility					0.70	0.031
VacA paralog	HP0609	Unknown					0.39	0.003
MdaB	HP0630	Modulator of drug activity	0.62	0.023	0.79	0.092		
FrxA	HP0642	Flavin oxidoreductase	0.58	<0.001	0.82	0.100		
GalU	HP0646	UDP-glucose pyrophosphorylase	0.76	<0.001	0.8	0.002		
Pfr	HP0653	Ferritin	0.65	0.004	0.97	0.940		
NrdA	HP0680	Ribonucleoside diphosphate reductase	0.71	0.027	0.74	0.055		
NusA	HP1514	Transcription termination factor	0.58	0.01	0.65	0.019		
PlpA	HP1564	Predicted OMP	0.72	0.046	0.84	0.280		
Upregulated[e]								
Hypothetical	HP0014	Unknown	1.63	0.004	1.34	0.054	2.21	0.017
Icd	HP0027	Isocitrate dehydrogenase	1.46	0.004	1.38	0.039		
DnaK	HP0109	Heat shock protein	1.31	0.042	1.04	0.760		
Hypothetical	HP0162	Unknown					1.40	0.028
PdgA	HP0310	Polysaccharide deacetylase	1.56	0.039	1.46	0.065		
CagA	HP0547	Virulence factor	2.86	<0.001	2.41	0.002	1.82	0.027
HefB	HP0606	Membrane fusion protein	1.31	<0.001	1.19	0.090		
Hypothetical	HP0688	Unknown	1.31	0.009	1.1	0.300		
GuaB	HP0829	IMP dehydrogenase	1.30	0.002	1.17	0.058		
RdxA	HP0954	Nitroreductase/metronidazole resistance	1.30	0.032	1.21	0.180		
FtsZ	HP0979	Cell division	1.40	0.019	1.23	0.049		
HopQ	HP1177	OMP	1.65	0.029	1.34	0.170	1.71	0.009
RocF	HP1399	Arginase	1.30	0.093	1.38	0.008		
Hypothetical	HP1588	Unknown	1.70	0.005	1.56	0.061	1.42	0.016

[a] The Swiss-Prot and NCBI nonredundant databases were searched.
[b] Corresponding gene numbers for the respective proteins in strain 26695.
[c] Known or predicted function of protein. OMP, outer membrane protein.
[d] *P* values were determined by the Student *t* test.
[e] Downregulated or upregulated in response to growth in a medium containing 1.25% NaCl, compared to growth in a medium containing either 0.25% or 0.5% NaCl.

were matched across the gel series and were analyzed to detect changes in expression (Fig. 1A to C). Protein features exhibiting statistically significant changes in expression in response to varying salt concentrations were identified by mass spectrometry (Table 2). Among these, CagA exhibited one of the greatest increases in expression in response to high-salt conditions (Table 2 and Fig. 1D). Proteomic analysis of the carcinogenic *H. pylori* strain 7.13 revealed a similar increase in CagA expression in response to high-salt conditions (Table 2). An outer membrane protein (HopQ) and two hypothetical proteins (HP0014 and HP1588) also exhibited the same pattern of regulation in both strains (Table 2).

Transcriptional analyses. To determine whether the changes observed in protein expression occurred at the transcriptional level, we analyzed the transcript levels of the relevant genes. For these experiments, we compared *H. pylori* 26695 grown in media containing two different salt concentrations (BB-FBS-0.5% and BB-FBS-1.25%) (Fig. 2). In agreement with the results obtained at the protein level, the transcript levels of *cagA*, *hopQ*, HP0014, *icd*, *pgdA*, HP0688, *ftsZ*, and HP1588 were significantly higher in bacteria grown in BB-FBS-1.25% than in bacteria grown in BB-FBS-0.5% (P, <0.05 by the t test). The transcript level for *mdaB* was significantly lower in bacteria grown in BB-FBS-1.25% than in bacteria grown in BB-FBS-0.5%. These results suggest that the differential expression of proteins shown in Table 2 can be attributed in most cases to alterations in gene transcription.

Analysis of salt-regulated CagA expression in Colombian *H. pylori* strains. In a previous study, we noted that the growth of some *H. pylori* strains (such as 26695) under high-salt conditions

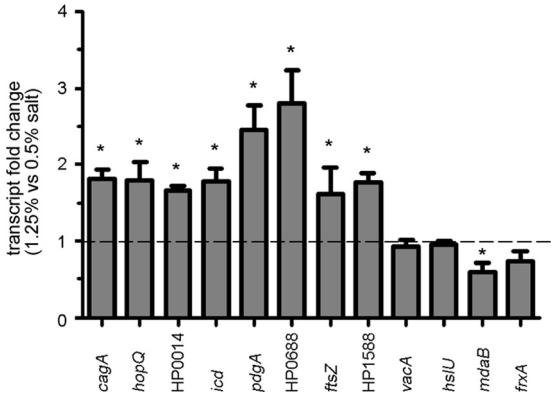

FIG 2 Real-time PCR analysis of transcripts corresponding to differentially expressed proteins identified in the proteomic analysis. *H. pylori* 26695 was cultured for 15 h in BB-FBS-0.5% or BB-FBS-1.25%. Bacterial RNA was isolated and processed, and real-time PCR was performed as described in Materials and Methods to assess the transcription of genes encoding proteins that were identified as differentially expressed, based on 2D-DIGE analysis. The transcript levels of each gene were normalized to the transcript levels of 16S rRNA. For each gene, normalized transcript values from bacteria grown under high salt-conditions (BB-FBS-1.25%) were compared to the corresponding values from bacteria grown under normal salt conditions (BB-FBS-0.5%) to yield a fold change value. The dashed line is provided as a reference and indicates a value of 1 (no change). The results shown (mean ± standard deviation) are from a single experiment with triplicate samples. The experiment was repeated 3 times with similar results. Asterisks indicate genes for which the transcript levels in bacteria cultured under high-salt conditions were significantly different from the transcript levels in bacteria cultured under normal salt conditions (P, <0.05 by Student's t test).

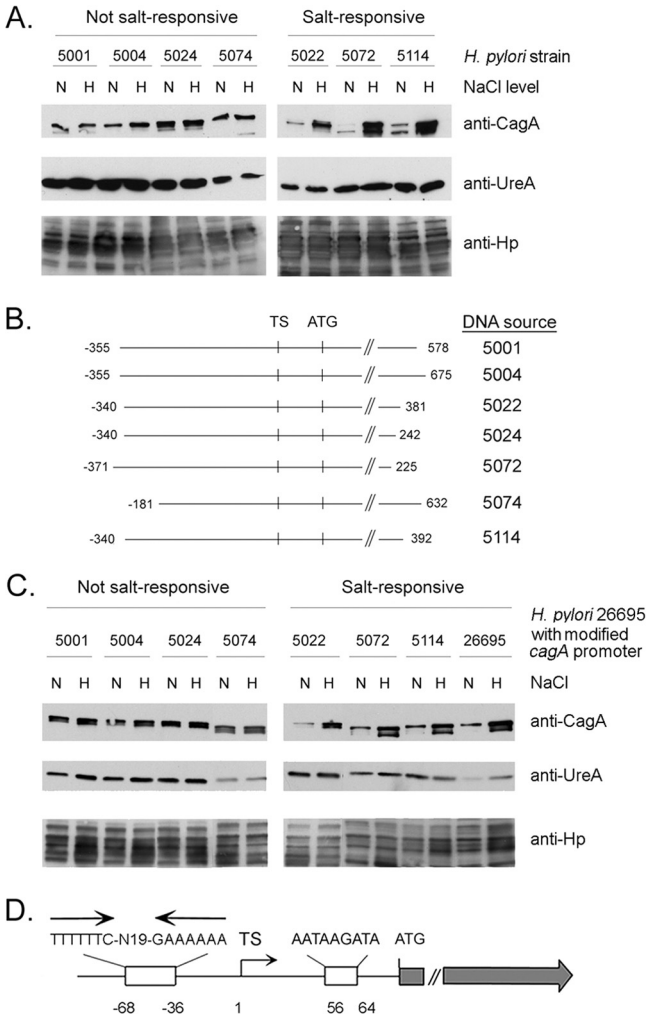

FIG 3 Heterogeneity among *H. pylori* strains in salt-regulated CagA expression. Salt-regulated expression of CagA was analyzed in *H. pylori* strains isolated from patients in regions of Colombia with a low risk (strains 5001, 5004, 5022, 5024) or a high risk (strains 5072, 5074, 5114) for gastric cancer. (A) *H. pylori* strains were cultured for 15 h in BB-FBS-0.5% (N, normal salt concentration) or BB-FBS-1.1% (H, high salt concentration). Cell extracts were standardized by protein concentration, and 5-μg aliquots were analyzed by SDS-PAGE and immunoblotting with an anti-CagA antibody (1:6,000). Consistent with previous results, CagA was visualized as either a single band or a doublet, and these bands were absent in *cagA* mutant strains (25). The basis for the appearance of a doublet in some cases is not known. Blots then were stripped and reprobed sequentially with an anti-UreA antiserum (1:10,000) and an anti-*H. pylori* antiserum (1:10,000). (B) *H. pylori* 26695 *cagA*::*catrdx*-9 was transformed with 1.2-kb DNA fragments (500 bp upstream and 700 bp downstream of the *cagA* transcriptional start site) derived from the indicated Colombian strains. The schematic illustrates chromosomal regions that were altered in the resulting transformants. Nucleotide numbers are relative to the *cagA* transcriptional start site and designate the chromosomal regions in *H. pylori* 26695 that were replaced with sequences derived from Colombian strains. Vertical lines indicate the *cagA* transcriptional start site (TS) and the *cagA* ATG translation initiation site. (C) The transformants schematized in panel B were cultured and analyzed for CagA expression as described for panel A. (D) Schematic illustrating an AT-rich inverse repeat in the *cagA* promoter region of *H. pylori* 26695, the *cagA* TS, an AATAAGATA motif associated with high basal CagA expression levels, and the *cagA* ATG translation initiation site.

resulted in increased CagA expression, whereas this phenomenon was not observed in other strains (25). To further investigate strain-specific differences in salt-regulated CagA expression, we analyzed a panel of 36 *H. pylori* Colombian strains that were isolated from patients in two regions in Colombia with differing incidences of gastric cancer (10). In these experiments, we found that some of the strains failed to grow in a medium containing 1.25% added NaCl, and therefore, experiments were conducted using a medium containing 1.1% added NaCl. There was marked variation in the ability of these strains to upregulate CagA expression in response to high-salt conditions. For example, CagA expression was upregulated when strains 5022, 5072, and 5114 were grown in BB-FBS-1.1% compared to expression when the same strains were grown in BB-FBS-0.5% (Fig. 3A). In contrast, little or no increase in CagA expression was observed when strains 5001, 5004, 5024, and 5074 were grown under high-salt conditions (Fig. 3A). To facilitate a quantitative analysis, we used densitometry to compare the levels of CagA to levels of a control protein (UreA), as described in Materials and Methods; a previous study showed that *ureA* expression does not change when *H. pylori* is grown under high-salt conditions (25). This approach allowed the calculation of a CagA induction value (a ratio of the level of CagA expression when a strain is grown under high-salt conditions to the level of CagA expression when the same strain is grown under low-salt conditions). Using this approach, we quantified salt-responsive changes in CagA expression among all 36 *H. pylori* Colombian strains. As shown in Table 3, 17 strains showed CagA induction values of >2.0, and 19 strains showed CagA induction values of ≤ 2.0.

Correlation between salt-regulated CagA expression and other bacterial or host characteristics. The hospitals where the

TABLE 3 CagA expression in response to salt in Colombian *H. pylori* strains[a]

Strain	CagA fold induction[b]	Risk region[c]	MLST result[d]	Basal CagA level[e]	Histopathology score[f]
5001	0.8 ± 0.2	Low	hpAfrica1	Low	2.3
5004	1.4 ± 0.2	Low	hpAfrica1	Low	2.3
5005	1.1 ± 0.3	Low	hpAfrica1	Low	1.7
5010	1.4 ± 0.6	Low	hpAfrica1	Low	2.0
5016	1.4 ± 0.2	Low	hpAfrica1	Low	3.5
5017	1.1 ± 0.3	Low	hpAfrica1	Low	3.5
5018	0.7 ± 0.2	Low	hpAfrica1	Low	2.3
5024	1.5 ± 0.2	Low	hpAfrica1	Low	2.0
5028	2.0 ± 0.6	Low	hpEurope	Low	3.8
5038	1.6 ± 0.4	Low	hpEurope	High	3.5
5051	1.4 ± 0.1	Low	hpAfrica1	High	2.0
5054	1.0 ± 0.3	Low	hpAfrica1	Low	2.3
5068	1.3 ± 0.1	High	hpEurope	High	4.3
5074	1.5 ± 0.1	High	hpEurope	High	4.7
5090	1.4 ± 0.2	High	hpEurope	Low	2.0
5093	1.9 ± 0.4	High	hpEurope	High	4.5
5096	1.3 ± 0.2	High	hpEurope	Low	4.3
5104	0.8 ± 0.1	High	hpEurope	High	3.5
5112	1.6 ± 0.2	High	hpEurope	High	4.3
5007	2.4 ± 0.5	Low	hpAfrica1	Low	2.0
5009	3.6 ± 0.8	Low	hpEurope	Low	1.7
5022	3.7 ± 0.7	Low	hpEurope	Low	4.5
5032	5.5 ± 0.9	Low	hpEurope	High	3.5
5043	2.2 ± 0.3	Low	hpEurope	High	5.3
5056	2.3 ± 0.7	High	hpEurope	High	4.5
5057	2.7 ± 0.3	High	hpEurope	High	3.5
5069	2.4 ± 0.1	High	hpEurope	High	4.3
5072	6.4 ± 2.1	High	hpEurope	Low	3.5
5080	2.2 ± 0.1	High	hpEurope	Low	3.5
5082	3.9 ± 1.4	High	hpEurope	High	4.8
5086	2.1 ± 0.5	High	hpEurope	High	4.6
5091	3.9 ± 0.8	High	hpEurope	Low	4.5
5095	3.4 ± 1.4	High	hpEurope	High	2.0
5097	4.4 ± 2.0	High	hpEurope	Low	4.4
5100	3.7 ± 1.3	High	hpEurope	High	2.3
5114	3.6 ± 0.1	High	hpEurope	Low	4.4

[a] The proportion of strains exhibiting CagA induction values of >2.0 was analyzed to detect possible correlations with the site of strain isolation (risk region), the MLST type, the basal level of CagA expression, and the histopathology score. A strong association was observed between the European MLST type and CagA induction values of >2.0 (P, 0.034 by the t test). P values (by the t test) of 0.054, 0.503, and 0.088 were obtained in comparisons of risk region, basal CagA level, and histopathology score, respectively, with the proportion of strains exhibiting CagA induction values of >2.0.

[b] CagA expression was analyzed by immunoblotting, CagA signal intensities were normalized to UreA signal intensities for each sample (i.e., CagA/UreA), and relative levels of expression (CagA fold induction) were quantified by comparing the normalized CagA signal intensity of each strain grown in 1.1% NaCl with normalized CagA signals for the same strain grown in 0.5% NaCl. For each strain, the CagA signal was calculated for a minimum of 3 independent biological samples, and the mean ± standard error is presented.

[c] *H. pylori* strains were isolated from regions of Colombia with either a high or a low risk for the development of gastric cancer.

[d] MLST analysis was used previously (10) to characterize the ancestral origins of the *H. pylori* strains. hpAfrica1 denotes isolates classified as either West African (hspWAfrica) or South African (hspSAfrica).

[e] Previous work had shown (23) that the Colombian strains used in this study demonstrated heterogeneity in the basal levels of CagA expressed when they were grown in a medium containing 0.5% NaCl. Here we classify strains expressing basal CagA levels, relative to strain 5001, of <2 or >2 as demonstrating low or high basal levels of CagA, respectively.

[f] Indicating the severity of lesions observed in the gastric tissue from which the *H. pylori* strains were isolated.

Colombian *H. pylori* strains were isolated are located in two cities with differing gastric cancer incidences: Túquerres in the Andes Mountains, where the gastric cancer incidence is high, and Tumaco on the coast, where the gastric cancer incidence is low. These Colombian *H. pylori* strains were previously classified as having either a predominantly African origin or a predominantly European origin, based on multilocus sequence type (MLST) analysis (10). When grown in BB-FBS-0.5%, *H. pylori* strains of European origin were more likely to express high basal levels of CagA, and strains expressing high basal levels of CagA were associated with more-advanced preneoplastic gastric lesions (23).

As shown in Fig. 4, salt-induced upregulation of CagA expression was detected more commonly in strains of European origin than in strains of African origin; this pattern was observed when we analyzed either European strains from the region of low gastric cancer risk ($P = 0.0016$) or European strains from the region of high gastric cancer risk ($P = 0.0034$) (Fig. 4B). The proportion of European strains (16 out of 25) showing CagA induction values of >2.0 was significantly higher than that observed for the African strains (1 out of 11) ($P = 0.0034$) (Table 3). The levels of salt-responsive CagA expression were not significantly different for strains from the Colombian regions of low versus high gastric cancer risk ($P = 0.0984$) (Fig. 4A); correspondingly, the proportions of strains showing CagA induction levels of >2.0 were not significantly different for strains from the low-risk region (5 of 17) and strains from the high-risk region (12 of 19) ($P = 0.0543$) (Table 3). Of the 5 strains from the low-risk region showing CagA induction values of >2.0, 4 were of European origin. Finally, salt-responsive CagA expression did not correlate with the basal levels of CagA expressed by each strain ($r^2 = 0.007$; $P = 0.6101$) (Table 3). No significant difference was observed when we compared CagA induction values of strains with lower basal levels (i.e., CagA levels of <2.0 relative to strain 5001) and higher basal levels (i.e., ≥2.0 relative to strain 5001) ($P = 0.7783$) of CagA (Fig. 4C). Consistent with this, the proportion of strains showing CagA induction levels of >2 among strains with lower basal levels of CagA (8 of 20) was not significantly different from that observed in strains with higher basal levels of CagA (9 of 16) ($P = 0.5027$) (Table 3). When we compared the levels of salt-induced CagA expression in 36 strains with the histopathology scores of the patients from whom these *H. pylori* strains were isolated, no significant association was observed ($r^2 = 0.055$; $P = 0.1671$) (Table 3). Thus, no significant difference was observed between the mean CagA induction values of strains from patients with precancerous gastric histology (histology score, ≥3.0, indicative of atrophic gastritis with or without intestinal metaplasia or dysplasia) and those of strains from patients with nonprecancerous gastric histology (histology score, <3.0, indicative of nonatrophic gastritis) ($P = 0.1032$) (Table 3).

The analyses discussed above did not account for differences among strains in the basal level of CagA expression, which was previously shown to be associated with the severity of gastric pathology (23). Therefore, we performed the same analyses on subgroups of strains previously classified as having either lower basal levels (expression level relative to strain 5001, <2.0) ($n = 20$) or higher basal levels (expression level relative to strain 5001, ≥2.0) ($n = 16$) of CagA expression when grown under normal salt conditions (23). As shown in Fig. 4D, among the 20 strains expressing low basal levels of CagA, there was a significant association between CagA induction values and the corresponding histopathol-

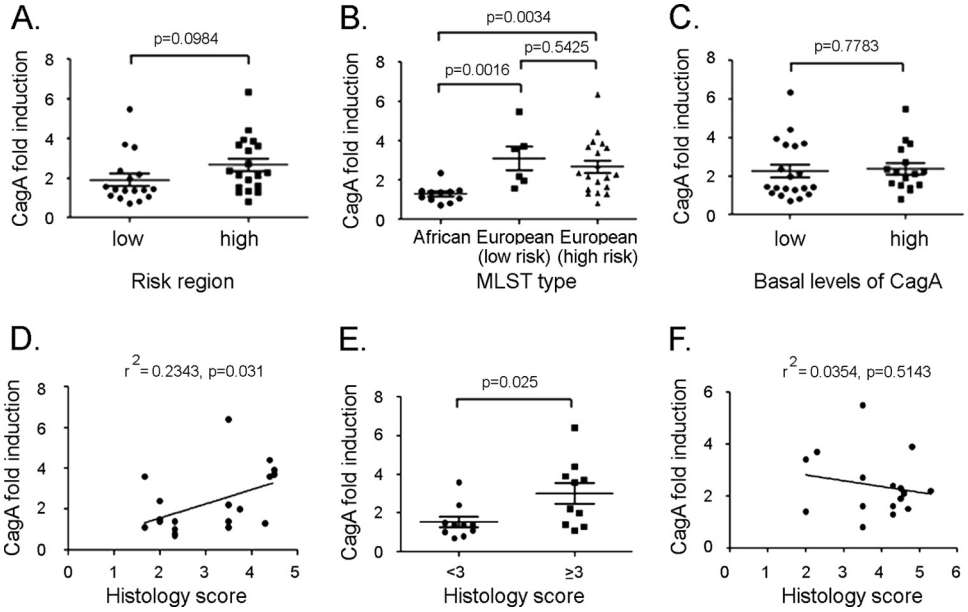

FIG 4 Relationships between salt-responsive CagA expression and site of *H. pylori* isolation, MLST type, basal level of CagA expression, and gastric histopathology score. This figure analyzes *H. pylori* strains isolated from patients living in regions of Colombia with either a low or a high risk for gastric cancer. *H. pylori* strains were cultured for 15 h in BB-FBS-0.5% or BB-FBS-1.1%. CagA expression in each strain was then analyzed by immunoblotting with an anti-CagA antibody and was quantified as described in Materials and Methods. For each strain, the effect of salt on CagA expression is expressed as a CagA induction value, which was calculated by comparing CagA expression levels under high-salt conditions (BB-FBS-1.1%) with those under lower-salt conditions (BB-FBS-0.5%). (A) Relationship between salt-responsive CagA expression and site of strain isolation (locations in Colombia with a high or a low risk for gastric cancer). (B) Relationship between salt-responsive CagA expression and the ancestral origin of the *H. pylori* strains, based on MLST analysis. The European strains in this analysis were subdivided according to whether they were isolated from regions of low or high cancer risk. (C) Relationship between salt-responsive CagA expression and basal levels of CagA expression (i.e., levels of CagA expressed when bacteria were cultured in BB-FBS-0.5%). Twenty strains had low basal levels of CagA expression (CagA expression levels, <2.0 compared to strain 5001), and 16 strains had higher basal levels of CagA expression (CagA expression levels, ≥2.0 relative to strain 5001) (23). (D, E, and F) Relationship between salt-responsive CagA expression and gastric histopathology score. The 20 strains analyzed in panels D and E all had low basal levels of CagA expression, and the 16 strains analyzed in panel F had higher basal levels of CagA expression, as described above for panel C. Panel E analyzes a relationship between gastric histology scores and salt-responsive CagA expression (histology scores of <3.0 correspond to nonatrophic gastritis [10 strains]; scores of ≥3.0 correspond to precancerous lesions [10 strains]). Mean fold induction values ± standard deviations are shown in panels A, B, C, and E. *P* values were determined by using Student's *t* test, and r^2 values were determined by linear regression analysis.

ogy scores ($r^2 = 0.2343$; $P = 0.031$). Consistent with the results in Fig. 4D, the CagA induction values of strains from patients with nonprecancerous lesions (histology score, <3.0) were significantly lower than those of strains from patients with precancerous lesions (histology score, ≥3.0) ($P = 0.025$), when this group of strains was analyzed (Fig. 4E). In contrast, in an analysis of 16 strains expressing high basal levels of CagA, there was no significant association between CagA induction values and histopathology scores (Fig. 4F) ($r^2 = 0.0354$; $P = 0.5143$).

Strain-specific sequences upstream of *cagA* influence salt-regulated CagA expression. We next sought to determine experimentally if the observed heterogeneity in salt-regulated CagA expression among the Colombian strains was attributable to differences in sequences upstream of the *cagA* ATG translation initiation site. For this purpose, we amplified 1.2-kb sequences (0.5 kb upstream and 0.7 kb downstream of the *cagA* transcriptional start site) from the 7 Colombian strains described above, and we introduced these sequences into *H. pylori* strain 26695 *cagA*::*catrdx*-9 (Fig. 3B). These experiments resulted in replacement of the *cat-rdx* cassette with a minimum of 181 bp upstream of the *cagA* transcriptional start site (or 285 bp upstream of the *cagA* ATG initiation site), derived from Colombian strains. The regions of the *H. pylori* 26695 genome replaced with sequences derived from Colombian strains are shown in Fig. 3B. As shown in Fig. 3C (left), *H. pylori* 26695 *cagA*::*catrdx*-9 transformants containing DNA derived from strain 5001, 5004, 5024, or 5074 showed little or no upregulation of CagA expression when grown in BB-FBS-1.1% compared to BB-FBS-0.5%. In contrast, *H. pylori* 26695 *cagA*::*catrdx*-9 transformants harboring DNA derived from strain 5022, 5072, or 5114 showed increased CagA expression when grown in a high-salt medium (Fig. 3C, right). Thus, the salt-responsive CagA expression phenotype of transformants correlated with the phenotype of the corresponding wild-type strains that were the sources of transformed DNA (compare Fig. 3A and C). These results indicate that variation among *H. pylori* strains in levels of CagA expression is attributable to variations in the region of DNA incorporated into *H. pylori* 26695 *cagA*::*catrdx*-9. The DNA incorporated into *H. pylori* 26695 *cagA*::*catrdx*-9 from Colombian strains contains the *cagA* transcriptional initiation site, an AT-rich inverted repeat important for *cagA* transcription (34), an AATAAGATA motif important for determining basal levels of CagA protein expression (23), and the ATG translation initiation site (Fig. 3D).

Comparative analysis of sequences upstream of *cagA*. In an effort to identify sequence differences that may account for the variation in salt-regulated CagA expression among *H. pylori* strains, we analyzed a ~1.2-kb region (0.5 kb upstream and 0.7 kb downstream of the transcriptional start site) that had been se-

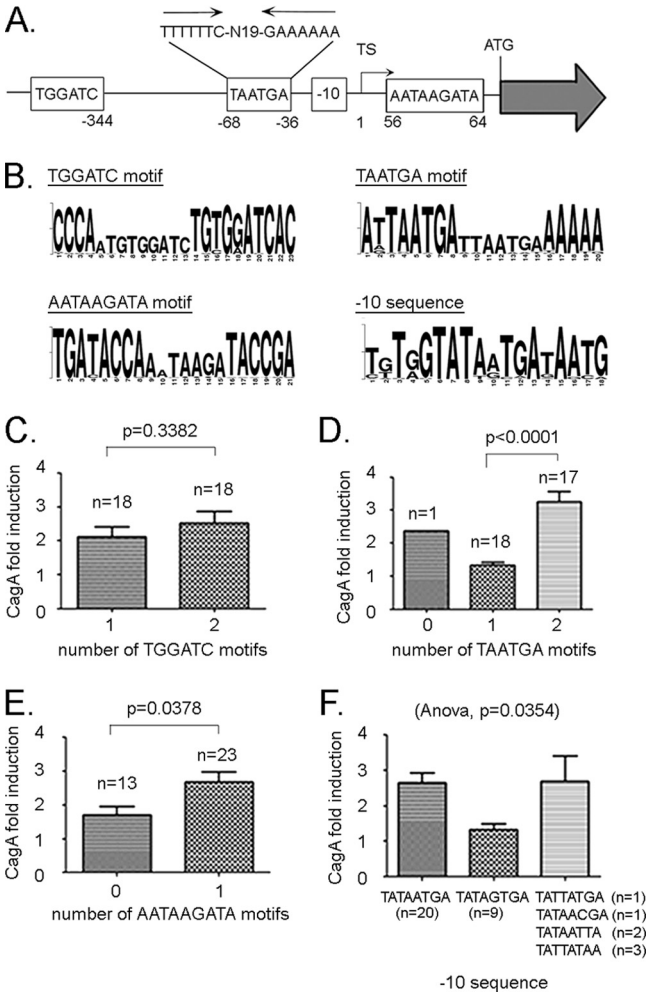

FIG 5 Relationship between salt-responsive CagA expression and strain-specific sequence variation in the region upstream of *cagA*. (A) Nucleotide sequence analysis of the region upstream of the *cagA* ATG initiation site in 36 Colombian *H. pylori* strains (23) revealed 4 regions that differed substantially among the strains. (B) Nucleotide sequence variation in these regions among the 36 strains is shown by WebLogo analysis. Each position in the alignment is represented by a stack of letters (nucleotides); the height of each letter is proportional to the observed frequency of each nucleotide. (C to F) Correlation of salt-responsive CagA expression with the number of TGGATC motifs (C), TAATGA motifs (D), or AATAAGATA motifs (E) and the type of −10 sequence (F). *P* values were calculated using the *t* test (C to E) or ANOVA (F). For each strain, salt-responsive CagA expression is shown as a "CagA fold induction value," calculated by comparing CagA expression levels under high-salt conditions (BB-FBS-1.1%) with those under lower-salt conditions (BB-FBS-0.5%).

quenced from all 36 strains previously. We previously identified several strain-specific differences upstream and downstream of the *cagA* transcriptional start site (Fig. 5A and B) (23). These include variation in the number of copies of a TGGATC motif (1 or 2 copies) located 344 bp upstream of the *cagA* transcriptional start site and variation in the number of copies of a TAATGA motif (0, 1, or 2 copies) located within an AT-rich region of the *cagA* promoter that contains an inverted repeat (Fig. 5A and B). The predicted −10 sequences of *cagA* promoters also differed among the strains and included TATAATGA, TATAGTGA, TATTATGA, and TATTATAA (Fig. 5A and B). In addition, there was variation among strains in the presence of a TGn motif located immediately upstream of the −10 promoter sequence (i.e., extended −10 sequence) (12). Finally, there was variation among strains in the presence of an AATAAGATA motif located approximately 56 nucleotides downstream of the *cagA* transcription start site (40 nucleotides upstream of the *cagA* ATG initiation codon) (Fig. 5A and B). The AATAAGATA motif is found more commonly in Colombian *H. pylori* strains with high levels of basal CagA expression than in strains with lower basal levels of CagA (23).

Figure 5 shows an analysis that correlates the presence of various motifs with the salt-responsive CagA expression phenotype. There was a strong association between CagA induction values and the number of TAATGA motifs (Fig. 5D). Strains containing 2 TAATGA motifs within the inverted-repeat motif region showed significantly higher CagA induction values than strains with only a single TAATGA motif ($P < 0.0001$). Fifteen (88.2%) of the 17 strains showing CagA induction values of >2.0 contained 2 TAATGA motifs. In contrast, among the 19 strains showing CagA induction values of ≤2.0, only 3 (15.7%) contained 2 TAATGA motifs ($P < 0.0001$). The nucleotide sequence variation in this region and the location of the TAATGA motifs relative to the transcriptional start site are illustrated for representative strains with CagA induction values of ≤2 (1 TAATGA motif for strains 5001, 5004, 5024, 5074) and CagA induction values of >2 (2 TAATGA motifs for strains 5022, 5072, 5114, 26695) (Fig. 6; see also Fig. S1 in the supplemental material).

An association was also observed when CagA induction values were correlated with the presence or absence of the AATAAGATA motif ($P = 0.0378$) (Fig. 5E) and the type of −10 sequence ($P = 0.0354$) (Fig. 5F). The AATAAGATA motif was present in 14 (82%) of the 17 strains with CagA induction values of >2.0 and in 9 (47%) of 19 strains with CagA induction values of ≤2 ($P = 0.0411$). No statistically significant association was observed between salt-regulated CagA expression and the presence of the TGGATC motif ($P = 0.3682$) (Fig. 5C) or the presence of an extended −10 sequence ($P = 0.074$) (data not shown).

Introduction of deletion mutations into the region upstream of the *cagA* open reading frame (ORF). Next we sought to experimentally delineate regions of DNA upstream of *cagA* that may be involved in salt-regulated CagA expression. Multiple DNA segments upstream of the *cagA* ATG initiation site were deleted (Fig. 7A), and *H. pylori* strains harboring these mutations were then analyzed to evaluate salt-responsive CagA expression. Three *H. pylori* strains (harboring deletions designated Δ6, Δ7, and Δ8) did not show increased expression of CagA when grown under high-salt conditions (Fig. 7B and C). In contrast, strains containing deletions Δ11, Δ12, Δ15, Δ16, Δ17, Δ18, and Δ19 retained the capacity to upregulate CagA expression in response to high-salt conditions (Fig. 7B and C). Strains containing deletions Δ6, Δ7, and Δ8 all lack sequences located 66 to 26 nucleotides upstream of the *cagA* transcriptional start site, and they also lack the TAATGA motifs (Fig. 7A). Deletants Δ11, Δ12, Δ15, Δ16, Δ17, Δ18, and Δ19 contain DNA deletions that are located upstream of the TAATGA motifs (Fig. 7A).

Targeted mutagenesis of the TAATGA and AATAGATA motifs. The deletion mutagenesis experiments described above support the notion that the TAATGA motifs are important for salt-regulated CagA expression. To experimentally test whether the number of copies or the nucleotide content of the TAATGA motifs influences salt-responsive CagA expression, site-specific mu-

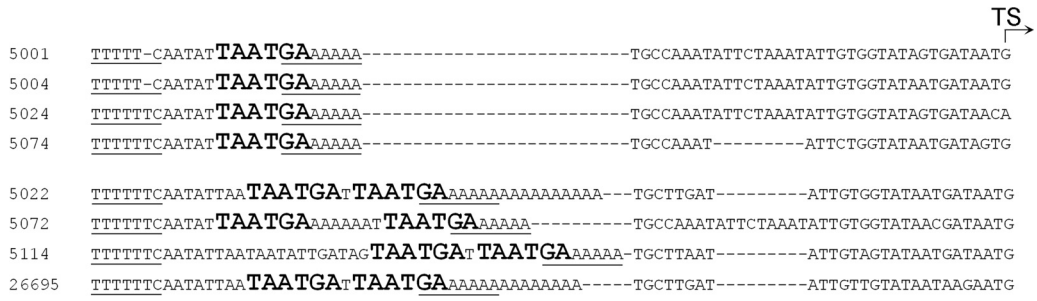

FIG 6 Strain-specific sequence variation in the *cagA* promoter region. This figure illustrates the TAATGA motif contained within an inverted repeat in the *cagA* promoter region. *H. pylori* strains 5022, 5072, 5114, and 26695 exhibited salt-regulated CagA expression, whereas strains 5001, 5004, 5024, and 5074 did not. The TAATGA motifs are in a larger font, and the nucleotides constituting the inverted repeat are underlined. The TAATGA motif is present in either 1 or 2 copies in the strains shown. The arrow indicates the *cagA* transcriptional start site (TS).

tations were introduced into this region of the *cagA* promoter. Two mutations (designated mut3 and mut21) each deleted one of the TAATGA motifs; one mutation (designated mut5) deleted both TAATGA motifs; and another mutation (designated mut22) altered the nucleotide sequence of the TAATGA motifs (Fig. 8A and B). As shown in Fig. 8C and D, *H. pylori* strains containing *cagA* promoters with mut3 ($P = 0.0044$), mut5 ($P = 0.0014$), mut21 ($P = 0.0017$), or mut22 ($P = 0.0032$) mutations did not significantly upregulate CagA expression in response to growth under high-salt conditions. These results indicate the importance of the TAATGA motifs in modulating increases in CagA expression in response to changes in salt concentration.

We also sought to determine whether the inverted repeat was required for salt-induced expression of CagA. For this purpose, we generated a mutant (mut23) in which the sequence comprising a portion of the inverted repeat was mutated. This mutant also exhibited a reduction in responsiveness to salt-induced upregulation of CagA expression (Fig. 8C and D).

Besides mutating the TAATGA motifs, we also analyzed the effects of mutations to the AATAAGATA motif (Fig. 8A and E) on salt-responsive CagA expression. As shown in Fig. 8D and F, strains harboring a deletion (mut16) or mutation (mut18) of the AATAAGATA motif demonstrated salt-responsive increases in CagA expression comparable to that observed in the wild-type strain (P, 0.5005 for mut16; P, 0.7762 for mut18). Thus, in contrast to the TAATGA motifs, the AATAAGATA motif does not play a role in the salt-responsive expression of CagA.

DISCUSSION

The risk of gastric cancer is determined by a combination of bacterial, host, and environmental factors. One of the environmental factors associated with an increased gastric cancer risk is the consumption of high levels of dietary salt (28, 36, 38). In Colombia, the consumption of high levels of salt (as measured by high urinary sodium-to-creatinine ratios) is associated with an increased risk for precancerous gastric lesions (chronic atrophic gastritis, intestinal metaplasia and dysplasia) compared to what is observed in persons who consume lower levels of salt (8). In some populations with a high incidence of gastric cancer (including China and Japan), dietary salt intakes have been reported to reach 55 g per day (19, 43). Based on an average adult stomach volume of 900 ml (15), gastric luminal salt concentrations of >300 mM may be expected in these populations.

In the current study, we used a global proteomic approach to analyze *H. pylori* strains cultured in media containing varying salt concentrations. Among the differentially expressed proteins identified, CagA exhibited the greatest increase in expression in response to high salt concentrations. We then examined salt-regulated CagA expression in 36 *H. pylori* strains isolated from patients in Colombia. Consistent with a previous report (25), salt-responsive CagA expression was not observed in all *H. pylori* strains. By analyzing DNA sequences upstream of the *cagA* ATG initiation site, we identified differences in the *cagA* promoter sequences of individual *H. pylori* strains. Strains demonstrating salt-induced increases in CagA expression were more likely to contain two copies of a TAATGA motif located within an AT-rich inverted-repeat sequence in the *cagA* promoter, while strains with little or no salt-responsive CagA expression were more likely to contain one copy of the TAATGA motif. Site-directed mutagenesis confirmed the importance of two copies of the TAATGA motif in mediating salt-responsive CagA expression. The presence of two copies of the TAATGA motif was found more frequently in *H. pylori* strains of European origin than in strains of African origin. Consistent with this, salt-responsive CagA expression was detected more commonly among strains of European origin than among strains of African origin.

The exact molecular mechanism by which the observed sequence variations lead to differences among strains in salt-responsive CagA expression is not yet known. As shown in Fig. 6, the TAATGA motif is located within an AT-rich inverted-repeat sequence upstream of the *cagA* genes. Tandem motifs are known to represent recognition elements for DNA binding by transcriptional regulators (27, 42). Therefore, one hypothesis is that this inverted-repeat region in the *cagA* promoter region serves as a binding site for a transcriptional regulator. However, examination of the inverted-repeat sequence using Mfold (RNA Institute, State University of New York at Albany) indicates that no nucleotide loop is predicted to be present. Another hypothesis is that tandem TAATGA motifs may function to ensure proper spacing between critical promoter elements (5, 17). Finally, it is possible that tandem copies of the TAATGA motif may serve a function analogous to that of a −35 element and thereby improve RpoD-RNA polymerase binding. Consensus −35 sequences have not been readily identifiable in *H. pylori* (12), but conserved AT-rich sequences have been detected in this region (39).

Several previous studies have analyzed the effects of varying salt concentrations on *H. pylori* gene expression (13, 14, 25) but there has been relatively little consensus in results among these

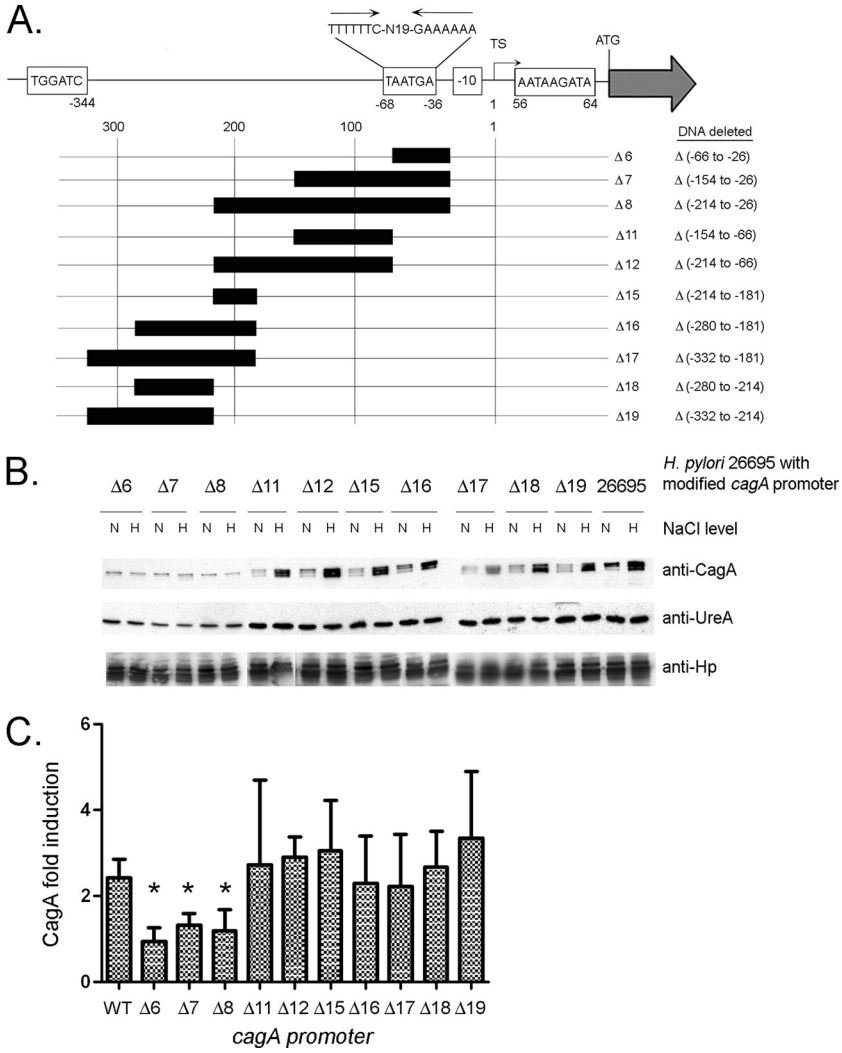

FIG 7 Analysis of CagA expression in strain 26695 mutants harboring deletions in the region upstream of the *cagA* transcriptional start site. (A) Cloned *cagA* fragments harboring deletions upstream of the *cagA* transcriptional start site were transformed into *H. pylori* 26695 *cagA*::*catrdx*-9. The filled rectangles represent the segments of DNA that were deleted. The locations of a TGGATC motif, a TAATGA motif, the *cagA* transcriptional start (TS), a AATAAGATA motif, and the *cagA* translation initiation site (ATG) are shown. (B) *H. pylori* 26695 *cagA*::*catrdx*-9 transformants harboring the indicated deletions in the *cagA* promoter region were cultured for 15 h in BB-FBS-0.5% (N, normal salt) or BB-FBS-1.1% (H, high salt). Lysates were generated as described in Materials and Methods and were standardized by protein concentration, and 5-μg aliquots were analyzed by Western blotting to detect CagA expression. (C) The levels of salt-induced CagA expression (CagA induction values) detected in the wild-type strain and mutants containing deletions in the *cagA* promoter region are shown. CagA induction values were calculated by comparing CagA expression levels under high-salt conditions (BB-FBS-1.1%) with those under lower-salt conditions (BB-FBS-0.5%). A minimum of 3 biological samples for each strain were analyzed. The mean CagA fold induction values ± standard errors are shown. The asterisks represent results that are significantly different from those for the wild-type strain (P, <0.005 by Student's t test).

studies. In the current study, in addition to detecting regulation of CagA expression, we detected changes in the expression of 30 other proteins in response to varying salt concentrations. The increase in CagA expression in response to salt was among the largest of the proteomic changes observed. Transcript levels of 6 of the 31 proteomic changes identified in this study (i.e., CagA, HopQ, HP0377, HP0630, HP1564, HP1588) were previously shown to be altered in response to salt, based on the use of array-based methodology (25), and salt responsive changes in *pfr* expression were detected using an RNase protection assay (14). Thus, despite an overall poor correlation between the list of salt-regulated genes identified in the current study and those identified in previous studies, we were able to corroborate several of the previous results. The relatively poor correlation among the results of multiple studies may be attributable to several factors, including limitations of earlier array-based methodology (and lack of validation of most array results), limitations in the ability of 2D-DIGE to resolve some proteins, and differences among studies in the choice of *H. pylori* strains selected for analysis.

We selected 12 of the differentially expressed proteins identified in the current study for further validation. Real-time PCR analysis showed that in the majority of cases (9 of 12 genes tested), salt-induced changes in transcript levels correlated with observed changes in protein expression. These included HopQ (an outer membrane protein) (24), PdgA (polysaccharide deacetylase) (33, 41), FtsZ (contributing cell division and cell morphology) (40),

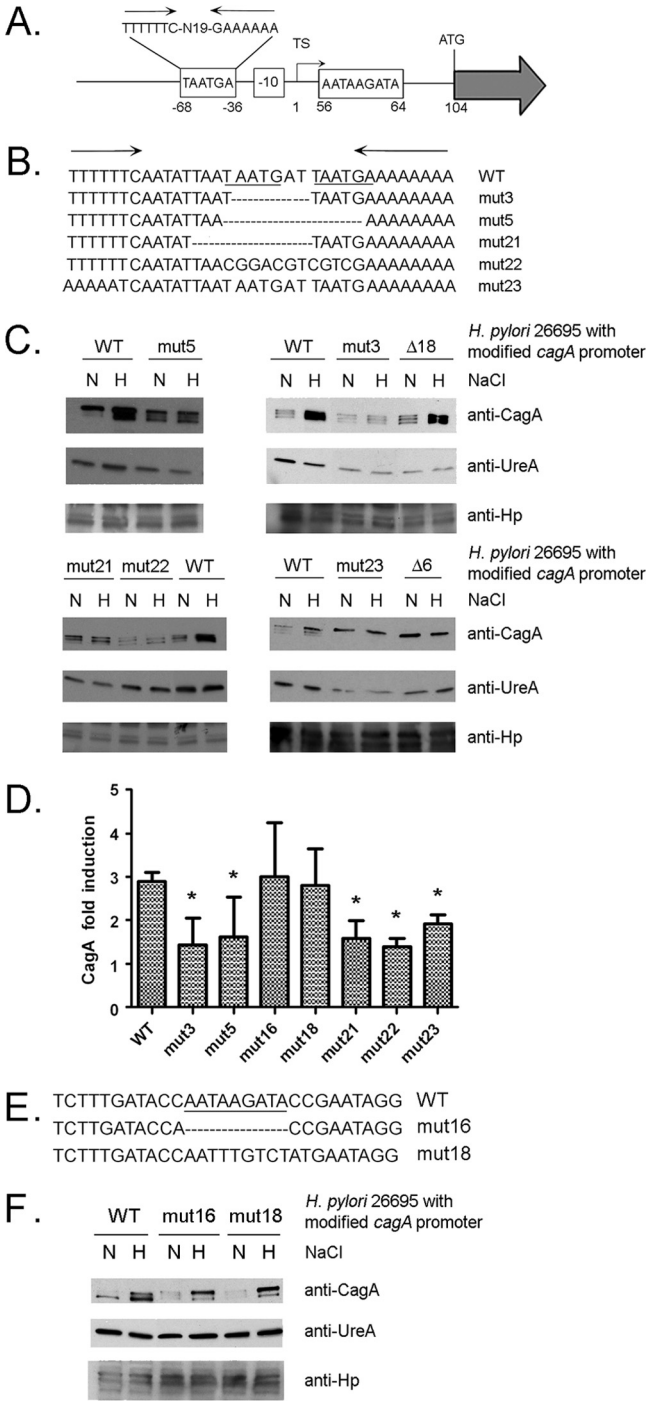

FIG 8 Identification of sequences required for salt-responsive CagA expression. This figure analyzes the effects of mutations to the TAATGA and AATAAGATA motifs on salt-responsive CagA expression. (A) Schematic illustrating the locations of the TAATGA and AATAAGATA motifs relative to the predicted *cagA* transcriptional start site (TS) and *cagA* ATG translation initiation site. (B) A cloned DNA segment containing 0.6 kb of DNA upstream and 0.6 kb downstream of the *cagA* transcriptional start site was mutated so that the indicated nucleotide changes (designated mut3, mut5, mut21, mut22, and mut23) were introduced into the inverted-repeat region (panel A) upstream of the *cagA* transcriptional start site. These mutations were then introduced into the corresponding chromosomal region upstream of *cagA* in *H. pylori* 26695 *cagA*::*catrdx*-9. (C) *H. pylori* mutants were cultured for 15 h in BB-FBS-0.5% (N, normal salt) or BB-FBS-1.1% (H, high salt). The bacteria were then harvested and lysed, and Western blotting was performed as described in Materials and Methods. (D) Levels of salt-induced CagA expression (CagA induction values) detected in the wild-type strain and mutants. CagA induction values were calculated by comparing CagA expression levels under high-salt conditions (BB-FBS-1.1%) with those under lower-salt conditions (BB-FBS-0.5%). Features of mut3, mut5, mut21, mut22, and mut23 are shown in panel B, and features of mut16 and mut18 are shown in panel E. The mean CagA induction value ± standard error was calculated based on a minimum of 6 biological samples for each strain. The asterisks represent results that are significantly different ($P, <0.005$ by Student's t test) from that observed for the wild-type strain. (E) Mutations were introduced to either delete (mut16) or mutate (mut18) the AATAAGATA motif found downstream of the *cagA* transcriptional start site (panel A). (F) Mutated plasmids harboring the mut16 and mut18 mutations were transformed into strain *H. pylori* 26695 *cagA*::*catrdx*-9, and transformants harboring the desired mutations were screened for salt-responsive CagA expression by Western blot analysis as described for panel C.

and multiple hypothetical proteins. The observed changes in the expression of FtsZ might account at least in part for alterations in *H. pylori* cell morphology that occur under high-salt conditions (13). At present, it is not clear whether the altered expression of these *H. pylori* proteins in response to salt is governed by a common mechanism. A search for motifs upstream of these genes, as well as in regions 300 bp upstream of other open reading frames in the *H. pylori* 26695 genome, did not reveal any identical matches to the tandem TAATGA motifs upstream of *cagA*.

CagA contributes to the pathogenesis of gastric cancer, and among CagA-positive *H. pylori* strains, sequence variation within CagA (such as the number and type of EPIYA motifs) (4, 18, 26) and variations in the basal level of CagA expression are associated with varying gastric cancer risks (23). In the current study, we examined whether salt-regulated CagA expression contributes to the risk of gastric cancer by correlating the salt-responsive phenotype of strains with the severity of gastric lesions in patients from whom the *H. pylori* strains were isolated. Among strains demonstrating relatively low basal levels of CagA (23), a correlation between the levels of salt-responsive CagA expression and the severity of gastric lesions was observed; less-severe gastric lesions were found in patients infected with strains exhibiting lower levels of salt-responsive CagA expression, while more-advanced gastric lesions were found in patients infected with strains exhibiting higher levels of salt-responsive CagA expression. In contrast, among strains with higher basal levels of CagA, no correlation between salt-responsive CagA expression and the severity of gastric lesions was observed. These findings suggest that in patients infected with strains that express low basal CagA levels, a high-salt diet can potentiate strain pathogenicity, whereas, in contrast, a high-salt diet may have relatively little effect on disease outcome in patients who are infected with strains that express high basal levels of CagA.

In conclusion, the present study demonstrates that exposure of *H. pylori* to varying environmental salt concentrations alters the expression of multiple proteins, among which CagA undergoes one of the greatest increases in expression in response to high-salt conditions. We show that there are strain-specific differences in the effect of salt on CagA expression, and we identify specific sequences upstream of *cagA* that account for this heterogeneity. Environmental factors such as a high-salt diet can potentially affect the disease outcome for *H. pylori*-infected persons through multiple mechanisms, and differences among strains in response to this environmental stimulus may be a relevant determinant of disease risk.

ACKNOWLEDGMENTS

This study was supported by NIH grants CA116087, AI068009, AI039657, CA77955, DK58587, and CA028842, Vanderbilt Digestive Disease Research Center grant DK058404, Vanderbilt Ingram Cancer Center, and the Department of Veterans Affairs.

REFERENCES

1. **Amieva MR, El-Omar EM.** 2008. Host-bacterial interactions in *Helicobacter pylori* infection. Gastroenterology **134**:306–323.
2. **Atherton JC, Blaser MJ.** 2009. Coadaptation of *Helicobacter pylori* and humans: ancient history, modern implications. J. Clin. Invest. **119**:2475–2487.
3. **Backert S, Tegtmeyer N, Selbach M.** 2010. The versatility of *Helicobacter pylori* CagA effector protein functions: the master key hypothesis. Helicobacter **15**:163–176.
4. **Basso D, et al.** 2008. Clinical relevance of *Helicobacter pylori cagA* and *vacA* gene polymorphisms. Gastroenterology **135**:91–99.
5. **Bellair M, Withey JH.** 2008. Flexibility of *Vibrio cholerae* ToxT in transcription activation of genes having altered promoter spacing. J. Bacteriol. **190**:7925–7931.
6. **Blaser MJ.** 2005. The biology of *cag* in the *Helicobacter pylori*-human interaction. Gastroenterology **128**:1512–1515.
7. **Blaser MJ, et al.** 1995. Infection with *Helicobacter pylori* strains possessing *cagA* is associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res. **55**:2111–2115.
8. **Chen VW, et al.** 1990. Risk factors of gastric precancerous lesions in a high-risk Colombian population. II. Nitrate and nitrite. Nutr. Cancer **13**:67–72.
9. **Cover TL, Blaser MJ.** 2009. *Helicobacter pylori* in health and disease. Gastroenterology **136**:1863–1873.
10. **de Sablet T, et al.** 2011. Phylogeographic origin of *Helicobacter pylori* is a determinant of gastric cancer risk. Gut **60**:1189–1195.
11. **Figueiredo C, et al.** 2002. *Helicobacter pylori* and interleukin 1 genotyping: an opportunity to identify high-risk individuals for gastric carcinoma. J. Natl. Cancer Inst. **94**:1680–1687.
12. **Forsyth MH, Cover TL.** 1999. Mutational analysis of the *vacA* promoter provides insight into gene transcription in *Helicobacter pylori*. J. Bacteriol. **181**:2261–2266.
13. **Gancz H, Jones KR, Merrell DS.** 2008. Sodium chloride affects *Helicobacter pylori* growth and gene expression. J. Bacteriol. **190**:4100–4105.
14. **Gancz H, Merrell DS.** 2011. The *Helicobacter pylori* ferric uptake regulator (Fur) is essential for growth under sodium chloride stress. J. Microbiol. **49**:294–298.
15. **Goetze O, et al.** 2007. The effect of macronutrients on gastric volume responses and gastric emptying in humans: a magnetic resonance imaging study. Am. J. Physiol. Gastrointest. Liver Physiol. **292**:G11–G17.
16. **Hatakeyama M.** 2004. Oncogenic mechanisms of the *Helicobacter pylori* CagA protein. Nat. Rev. Cancer **4**:688–694.
17. **Hidalgo E, Demple B.** 1997. Spacing of promoter elements regulates the basal expression of the *soxS* gene and converts SoxR from a transcriptional activator into a repressor. EMBO J. **16**:1056–1065.
18. **Higashi H, et al.** 2005. EPIYA motif is a membrane-targeting signal of *Helicobacter pylori* virulence factor CagA in mammalian cells. J. Biol. Chem. **280**:23130–23137.
19. **Howson CP, Hiyama T, Wynder EL.** 1986. The decline in gastric cancer: epidemiology of an unplanned triumph. Epidemiol. Rev. **8**:1–27.
20. **Huang JQ, Zheng GF, Sumanac K, Irvine EJ, Hunt RH.** 2003. Meta-analysis of the relationship between *cagA* seropositivity and gastric cancer. Gastroenterology **125**:1636–1644.
21. **Jones KR, Whitmire JM, Merrell DS.** 2010. A tale of two toxins: *Helicobacter pylori* CagA and VacA modulate host pathways that impact disease. Front. Microbiol. **1**:115. doi:10.3389/fmicb.2010.00115.
22. **Loh JT, Gupta SS, Friedman DB, Krezel AM, Cover TL.** 2010. Analysis of protein expression regulated by the *Helicobacter pylori* ArsRS two-component signal transduction system. J. Bacteriol. **192**:2034–2043.
23. **Loh JT, et al.** 2011. Analysis of *cagA* in *Helicobacter pylori* strains from Colombian populations with contrasting gastric cancer risk reveals a biomarker for disease severity. Cancer Epidemiol. Biomarkers Prev. **20**:2237–2249.
24. **Loh JT, Torres VJ, Algood HM, McClain MS, Cover TL.** 2008. *Helicobacter pylori* HopQ outer membrane protein attenuates bacterial adherence to gastric epithelial cells. FEMS Microbiol. Lett. **289**:53–58.
25. **Loh JT, Torres VJ, Cover TL.** 2007. Regulation of *Helicobacter pylori cagA* expression in response to salt. Cancer Res. **67**:4709–4715.
26. **Naito M, et al.** 2006. Influence of EPIYA-repeat polymorphism on the phosphorylation-dependent biological activity of *Helicobacter pylori* CagA. Gastroenterology **130**:1181–1190.
27. **Ni X, Westpheling J.** 1997. Direct repeat sequences in the *Streptomyces* chitinase-63 promoter direct both glucose repression and chitin induction. Proc. Natl. Acad. Sci. U. S. A. **94**:13116–13121.
28. **Nozaki K, et al.** 2002. Synergistic promoting effects of *Helicobacter pylori* infection and high-salt diet on gastric carcinogenesis in Mongolian gerbils. Jpn. J. Cancer Res. **93**:1083–1089.
29. **Parsonnet J, Friedman GD, Orentreich N, Vogelman H.** 1997. Risk for gastric cancer in people with CagA positive or CagA negative *Helicobacter pylori* infection. Gut **40**:297–301.
30. **Plummer M, et al.** 2007. *Helicobacter pylori* cytotoxin-associated genotype gastric precancerous lesions. J. Natl. Cancer Inst. **99**:1328–1334.
31. **Polk DB, Peek RM, Jr.** 2010. *Helicobacter pylori*: gastric cancer and beyond. Nat. Rev. Cancer **10**:403–414.
32. **Shaffer CL, et al.** 2011. *Helicobacter pylori* exploits a unique repertoire of type IV secretion system components for pilus assembly at the bacteria-host cell interface. PLoS Pathog. **7**:e1002237. doi:10.1371/journal.ppat.1002237.
33. **Shaik MM, Cendron L, Percudani R, Zanotti G.** 2011. The structure of *Helicobacter pylori* HP0310 reveals an atypical peptidoglycan deacetylase. PLoS One **6**:e19207. doi:10.1371/journal.pone.0019207.
34. **Spohn G, Beier D, Rappuoli R, Scarlato V.** 1997. Transcriptional analysis of the divergent *cagAB* genes encoded by the pathogenicity island of *Helicobacter pylori*. Mol. Microbiol. **26**:361–372.
35. **Suerbaum S, Michetti P.** 2002. *Helicobacter pylori* infection. N. Engl. J. Med. **347**:1175–1186.
36. **Tatematsu M, Nozaki K, Tsukamoto T.** 2003. *Helicobacter pylori* infection and gastric carcinogenesis in animal models. Gastric Cancer. **6**:1–7.
37. **Tegtmeyer N, Wessler S, Backert S.** 2011. Role of the *cag*-pathogenicity island encoded type IV secretion system in *Helicobacter pylori* pathogenesis. FEBS J. **278**:1190–1202.
38. **Tsugane S.** 2005. Salt, salted food intake, and risk of gastric cancer: epidemiologic evidence. Cancer Sci. **96**:1–6.
39. **Vanet A, Marsan L, Labigne A, Sagot MF.** 2000. Inferring regulatory elements from a whole genome. An analysis of *Helicobacter pylori* σ^{80} family of promoter signals. J. Mol. Biol. **297**:335–353.
40. **Varma A, Young KD.** 2004. FtsZ collaborates with penicillin binding proteins to generate bacterial cell shape in *Escherichia coli*. J. Bacteriol. **186**:6768–6774.
41. **Wang G, et al.** 2010. Peptidoglycan deacetylation in *Helicobacter pylori* contributes to bacterial survival by mitigating host immune responses. Infect. Immun. **78**:4660–4666.
42. **Withey JH, DiRita VJ.** 2006. The toxbox: specific DNA sequence requirements for activation of *Vibrio cholerae* virulence genes by ToxT. Mol. Microbiol. **59**:1779–1789.
43. **You WC, et al.** 1988. Diet and high risk of stomach cancer in Shandong, China. Cancer Res. **48**:3518–3523.